(19)日本国特許庁（ＪＰ）　(12)公開特許公報（Ａ）　(11)特許出願公開番号

特開平8－314779

(43)公開日　平成８年(1996)11月29日

(51)Int.Cl.[6]		識別記号	庁内整理番号	ＦＩ		技術表示箇所
Ｇ０６Ｆ	12/00	５１４	7623－5B	Ｇ０６Ｆ	12/00	５１４Ｍ
	13/00	３５７	7368－5E		13/00	３５７Ｚ
// Ｇ０６Ｆ	12/08		7623－5B		12/08	Ｗ
		３２０	7623－5B			３２０

審査請求　未請求　請求項の数４　ＯＬ　（全６頁）

(21)出願番号　特願平7－116826

(22)出願日　平成７年(1995)５月16日

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地

(72)発明者　榎本　俊一
愛知県尾張旭市晴丘町池上１番地株式会社日立製作所オフィスシステム事業部内

(72)発明者　木下　登
愛知県尾張旭市晴丘町池上１番地株式会社日立製作所オフィスシステム事業部内

(74)代理人　弁理士　小川　勝男

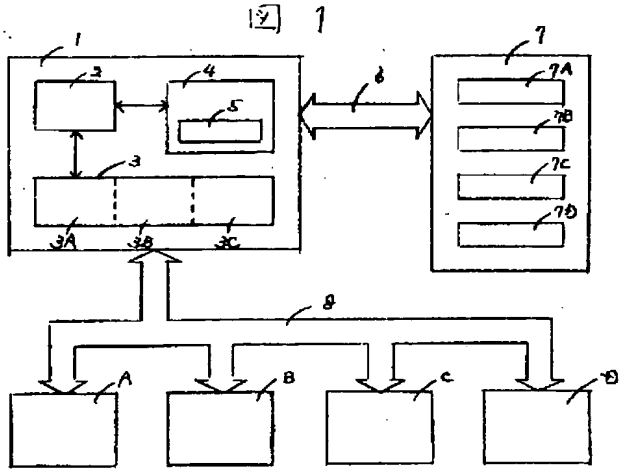
(54)【発明の名称】　サーバーシステム

(57)【要約】

【目的】特定のファイルアクセス要求元クライアントとキャッシュ間のファイルアクセスの処理速度を向上させ、システム全体の処理能力を向上でき、資源の効率運用が可能なサーバシステムを提供する。

【構成】本発明ではサーバーシステムのキャッシュを分割する。キャッシュの分割は、ファイルアクセス要求を発生させたファイルアクセス要求元毎に行ない、分割したキャッシュをキャッシュアロケーションテーブルの管理情報によって管理する。キャッシュアロケーションテーブルには、分割したキャッシュの使用状況を管理する情報とそのキャッシュの詳細情報を与える。ファイルアクセス要求元は、それぞれの分割したキャッシュを専用のキャッシュ領域として使用する。

図 1



1

【特許請求の範囲】

【請求項１】ハードディスクなどの外部記憶装置と、該外部記憶装置から読み出したデータを蓄積するキャッシュとを備え、前記外部記憶装置に格納されるファイルを使用する複数のクライアントが接続されるサーバーシステムにおいて、

前記キャッシュの領域を複数に分割し、該キャッシュの分割領域と該分割領域を割り当てたファイルアクセス要求元のクライアントとの対応を記憶する管理テーブルを備え、前記キャッシュの分割領域をクライアント毎に割り当てて使用することを特徴とするサーバーシステム。

【請求項２】請求項１記載のサーバーシステムにおいて、

一のクライアントのアクセス要求を受けたときに、前記管理テーブルを参照し、当該クライアントに割り付けた分割領域があれば、該分割領域を用い、なければ、空いている分割領域に当該クライアントを割り付けることを特徴とするサーバーシステム。

【請求項３】請求項２記載のサーバーシステムにおいて、

前記複数のクライアントに前記キャッシュを使用する優先順位を付け、

一のクライアントのアクセス要求を受けたときに、既に当該クライアントに割り付けた領域も、空き領域も無い場合に、優先順位を比較し、当該クライアントより優先順位の低いクライアントがあったら、該分割領域の専有クライアントを変更することを特徴とするサーバーシステム。

【請求項４】請求項３記載のサーバーシステムにおいて、更に、アクセス要求を出した前記クライアントより優先順位の低いクライアントが無かったら前記クライアントにはキャッシュを割り付けないことを特徴とするサーバーシステム。

【発明の詳細な説明】

【０００１】

【産業上の利用分野】本発明は、コンピュータシステムにおけるデータ転送処理方式に関し、特にサーバーシステムにおけるサーバーのキャッシュ分割管理に関する。

【０００２】

【従来の技術】一般に、クライアント／サーバーシステムにおいて、クライアントがサーバーシステムにアクセスし、サーバーシステムに接続された外部記憶装置のデータを転送する場合、外部記憶装置からのデータは、処理速度を向上させるため入出力バスを介してサーバーシステム内のキャッシュに読み込まれる。キャッシュに読み込まれたデータは、サーバーシステムに接続されたネットワークを介してクライアントに転送される。クライアントは、サーバーシステムより転送されたデータを処理し、必要であれば処理結果を再度サーバーシステムにアクセスし、外部記憶装置へ格納する。

2

【０００３】その後同一のクライアントが同じデータをアクセスする場合は、サーバーはサーバーシステム内のキャッシュデータを転送することにより、外部記憶装置のアクセスなしに高速にデータ転送処理を可能とする。

【０００４】しかし、同一のクライアントがキャッシュ内のデータを再度アクセスする前に、別のクライアントがキャッシュ内のデータとは異なるデータをサーバーにアクセスした場合、前のクライアントがキャッシュに蓄積したデータは、後のクライアントが外部記憶装置から読みだすデータに上書きされ、キャッシュ内のデータは前のクライアントが蓄積したデータとは異なってしまう。

【０００５】したがって、この状態において前のクライアントが前回と同じデータを再度アクセスした場合、サーバーは外部記憶装置へのアクセスが必要となり、キャッシュのもつ機能を有効に使用することができない。

【０００６】他の装置にキャッシュ内のデータを上書きされることなく書き込み／読み出し処理を行なう従来技術としては、特開平３－２８６３８２号公報に示されるメモリ制御装置のように複数のページバッファと複数のコントローラ間に任意の接続を行なうスイッチ回路を設け、システムの処理能力を向上させる技術があるが、複数のページバッファ、複数のコントローラ、その間の任意の接続を行なうスイッチ回路等が新たに必要となり、高価なシステムとなる。

【０００７】

【発明が解決しようとする課題】クライアント／サーバーシステムでは、複数のクライアントがサーバーに対してサービスを要求し、サーバーはサービスを要求したそれぞれのクライアントに対して特定のサービスを提供する。一般的には、サービスを要求するクライアントは、それぞれが別々の目的を持ってサーバーをアクセスするため、各クライアント毎に異なったファイル情報を要求する。つまり、キャッシュを備えたサーバーシステムにおいて、あるクライアントがサーバーのキャッシュとデータ処理中に他のクライアントがサーバーをアクセスした場合、サーバー内のキャッシュデータは、他のクライアントが要求するファイル情報に上書きされる。このため、クライアント側から見たサーバーのキャッシュ機能は有効に使用することができない。

【０００８】また、キャッシュデータが上書きされることなく書き込み／読み出し処理を行なうには、特開平３－２８６３８２号公報の技術のように、複数のページバッファを必要とする。さらに、複数のページバッファと複数のコントローラ間の任意の接続には、そのスイッチ回路が新たに必要であった。

【０００９】本発明の目的は、特別なスイッチ回路を用いることなく、キャッシュメモリ資源等を効率良く運用することができるサーバーシステムを提供することに有り、更に、キャッシュを分割管理することにより、キャ

ッシュデータの無効な上書きをなくし、特定のファイルアクセス要求元とキャッシュ間のファイルアクセスの処理速度を向上させ、システム全体の処理能力を向上できるサーバーシステムを提供することにある。

【００１０】

【課題を解決するための手段】前述した目的を達成するために、ハードディスクなどの外部記憶装置と、該外部記憶装置から読み出したデータを蓄積するキャッシュとを備え、前記外部記憶装置に格納されるファイルを使用する複数のクライアントが接続されるサーバーシステムにおいて、前記キャッシュの領域を複数に分割し、該キャッシュの分割領域と該分割領域を割り当てたファイルアクセス要求元のクライアントとの対応を記憶する管理テーブルを備え、前記キャッシュの分割領域をクライアント毎に割り当てて使用する。

【００１１】また、一のクライアントのアクセス要求を受けたときに、前記管理テーブルを参照し、当該クライアントに割り付けた分割領域があれば、該分割領域を用い、なければ、空いている分割領域に当該クライアントを割り付ける。

【００１２】更に、前記複数のクライアントに前記キャッシュを使用する優先順位を付け、一のクライアントのアクセス要求を受けたときに、既に当該クライアントに割り付けた領域も、空き領域も無い場合に、優先順位を比較し、当該クライアントより優先順位の低いクライアントがあったら、該分割領域の専有クライアントを変更し、アクセス要求を出した前記クライアントより優先順位の低いクライアントが無かったら前記クライアントにはキャッシュを割り付けない。

【００１３】即ち、本発明ではサーバーシステムのキャッシュを分割する。キャッシュの分割は、ファイルアクセス要求を発生させたファイルアクセス要求元毎に行ない、分割したキャッシュをキャッシュアロケーションテーブルの管理情報によって管理する。キャッシュアロケーションテーブルには、分割したキャッシュの使用状況を管理する情報とそのキャッシュの詳細情報を与える。ファイルアクセス要求元は、それぞれの分割したキャッシュを専用のキャッシュ領域として使用する。このようにして、ファイルアクセス要求元が蓄積したキャッシュ内のデータは、他のファイルアクセス要求元に上書きされることがなく、特定のファイルアクセス要求元とサーバーシステム内のキャッシュ間でのファイルアクセスの処理速度を向上させ、システム全体の処理能力を向上することができる。

【００１４】

【作用】本発明によれば、分割したキャッシュ領域とクライアント毎に割り当てるので、キャッシュ内のデータが他のクライアントからアクセス要求されたファイルによって上書きされることが無くなる。

【００１５】また、クライアントからアクセス要求を受

けたときに、分割領域を動的に割り当て、複数のクライアントにキャッシュを使用する優先順位を付けて管理するので、キャッシュ資源を有効に活用することができる。

【００１６】このように、キャッシュの分割領域がクライアント毎に個別に割り当てられるので、ハードディスクなどの外部記憶装置とサーバーシステム間のアクセスを減らすことが可能となり、サーバーシステムとファイルアクセス要求元との間の高速なファイルアクセスが可能となる資源利用効率の高いサーバーシステムが提供できる。

【００１７】

【実施例】第１図は、本発明が適用されるクライアント／サーバーシステムの概要を示す構成図である。サーバーシステム１は、サーバー機能を制御するＣＰＵ２と外部記憶装置７から読み出したデータを蓄積する分割可能なキャッシュ３とキャッシュ３領域の使用状況を管理するキャッシュアロケーションテーブル５（以下、テーブルと略称する）を記録するメモリ４を備えており、入出力バス６により外部記憶装置７と接続されている。さらに、サーバーシステム１は、ネットワーク８を介してクライアントＡ～Ｄに接続されている。以下、キャッシュ３をキャッシュ３Ａ、キャッシュ３Ｂ、キャッシュ３Ｃの３面に分割し、外部記憶装置７にデータ７Ａ、データ７Ｂ、データ７Ｃ、データ７Ｄが格納されている場合を例にとって説明する。

【００１８】本実施例では、キャッシュ分割方法として、クライアントＩＤを認識することによるキャッシュ分割を例にとって述べるが、サーバーシステムへのクライアント登録時の情報など、本発明はクライアントの認識手段に限定されるものではない。また、クライアントＡ～ＤのＩＤによるキャッシュ３分割時のキャッシュ割り当て優先順位は、Ａを最優先順位とし、以下Ｂ、Ｃ、Ｄの順に続く。

【００１９】第２図は、本発明のキャッシュ３管理に必要なキャッシュアロケーションテーブル５の一例である。本実施例において、テーブル５は分割キャッシュ領域の使用状態をビットで管理するテーブル５Ａ（以下、ビットマップ５Ａと略称する）とビットマップ５Ａの各ビット５Ａ１～５Ａ６に対応した分割キャッシュ領域の詳細情報５Ｂからなる。ビットマップ５Ａのビット数は、キャッシュ分割時の最大分割数を示し、本実施例においては６である。また、ビットマップ５Ａの各ビット５Ａ１～５Ａ６は、対応する分割キャッシュ領域が使用中であれば、ビットが１、なければ０になる。

【００２０】分割キャッシュ領域の詳細情報５Ｂは、ビットマップ５Ａの各ビット５Ａ１～５Ａ６毎に、分割キャッシュを使用しているクライアントのクライアントＩＤ５Ｂ１、分割キャッシュの先頭ブロックアドレス５Ｂ２と最終ブロックアドレス５Ｂ３、分割キャッシュのレ

ングス５Ｂ４、分割キャッシュに登録されているファイル名やディスクアドレス情報等が記録されたファイル情報５Ｂ５からなる。本実施例では、分割したキャッシュ３Ａ、３Ｂ、３Ｃに対応するビットマップ５Ａの各ビットは、５Ａ１～５Ａ２、５Ａ３～５Ａ４、５Ａ５～５Ａ６である場合を例にとって説明する。

【００２１】クライアントＡがデータ７Ａをアクセスするためには、クライアント／サーバーシステム間での所定のネットワーク手順に従って、クライアントＡはネットワーク８にクライアントＡのＩＤを送信し、他クライアントとの調停後、サーバーシステム１のアクセス権利を得る。クライアントＡから送信したＩＤは、サーバーシステム１が受信する。受信したＩＤはＣＰＵ２がクライアントＡのものであると認識する。ＩＤを認識したＣＰＵ２は、サーバーシステム１内のキャッシュ３の使用状況をテーブル５のビットマップ５Ａにより判断する。ＣＰＵ２はビットマップ５Ａのビット５Ａ１、５Ａ２が共に０で、キャッシュ３Ａに空きが有ると判断すると、キャッシュ３Ａに対応したビットマップ５Ａのビット５Ａ１と５Ａ２を０から１とする。また、ＣＰＵ２は、キャッシュ３Ａ領域の詳細情報５Ｂとして、ビットマップ５Ａのビット５Ａ１と５Ａ２に対応したクライアントＩＤ５Ｂ１、先頭ブロックアドレス５Ｂ２、最終ブロックアドレス５Ｂ３、レングス５Ｂ４、ファイル情報５Ｂ５を記録し、クライアントＡにキャッシュ３Ａを割り当てる。割り当てたキャッシュ３Ａには、外部記憶装置７から読みだしたデータ７Ａを転送、蓄積し、クライアントＡは、キャッシュ３Ａ内のデータ７Ａをアクセスする。

【００２２】次にクライアントＡがキャッシュ３Ａ内のデータ７Ａを処理中に、クライアントＢがデータ７Ｂをアクセスする。クライアントＢはクライアントＡと同様の手順にてクライアントＢのＩＤをネットワーク８に送信し、サーバーシステム１のアクセス権利を得る。受信したＩＤがクライアントＢであると認識したＣＰＵ２は、キャッシュ３の使用状況をテーブル５のビットマップ５Ａにより判断する。ＣＰＵ２はビットマップ５Ａのビット５Ａ１、５Ａ２は１であるがビット５Ａ３と５Ａ４が０で、キャッシュ３Ｂに空きがあると判断すると、キャッシュ３Ｂ領域に対応したビットマップ５Ａのビット５Ａ３と５Ａ４を０から１とする。また、ＣＰＵ２は、キャッシュ３Ｂ領域の詳細情報５Ｂとして、ビットマップ５Ａのビット５Ａ３と５Ａ４に対応したクライアントＩＤ５Ｂ１、先頭ブロックアドレス５Ｂ２、最終ブロックアドレス５Ｂ３、レングス５Ｂ４、ファイル情報５Ｂ５を記録し、クライアントＢにキャッシュ３Ｂを割り当てる。割り当てたキャッシュ３Ｂには、外部記憶装置７から読みだしたデータ７Ｂを転送、蓄積し、クライアントＢは、キャッシュ３Ｂ内のデータ７Ｂをアクセスする。また、この時、クライアントＡ用に割り当てたキャッシュ３Ａ内のデータ７Ａは、ＣＰＵ２のキャッシュ

分割制御によりクライアントＢのアクセスで消去されることはない。逆にクライアントＢ用に割り当てたキャッシュ３Ｂ内のデータ７Ｂも、ＣＰＵ２のキャッシュ分割制御によりクライアントＡのアクセスで消去されることはない。もちろん、他の優先順位の低いクライアントのアクセスにおいてもキャッシュ３Ａ、３Ｂ内のデータ７Ａ、７Ｂは消去されない。

【００２３】次に上記状態から同様の手順にてキャッシュ３Ｃをクライアント Ｃ用に割り当て、クライアントＡ、Ｂ、Ｃがそれぞれキャッシュ３Ａ、３Ｂ、３Ｃを使用中でキャッシュ３には他のクライアントのためにキャッシュを割り当てる領域がない場合に、クライアントＤがデータＤをアクセスする。クライアントＤはクライアントＡと同様の手順にてクライアントＤのＩＤをネットワーク８に送信し、サーバーシステム１のアクセス権利を得る。受信したＩＤがクライアントＤであると認識したＣＰＵ２は、キャッシュ３の使用状況をテーブル５のビットマップ５Ａにより判断する。ＣＰＵ２はビットマップ５Ａの各ビット５Ａ１～５Ａ６が１でキャッシュ３に空きがないと判断すると、テーブル５の分割キャッシュの詳細情報５Ｂからクライアント ＩＤ５Ｂ１を読み取り、キャッシュ３を使用中であるクライアントＡ、Ｂ、Ｃの中で最も分割時の優先順位が低いクライアントＣとアクセス要求元であるクライアントＤのキャッシュ分割時の優先順位をＩＤにより比較する。キャッシュ分割時の優先順位がクライアントＣよりクライアントＤが低いと判断したＣＰＵ２は、クライアントＤ用にキャッシュを割り当てることなく、クライアントＤは外部記憶装置７から読みだしたデータ７Ｄを直接アクセスする。また、テーブル５の内容も書き替えない。従って、クライアントＤによるアクセスでキャッシュ３内のデータが上書きされることはない。

【００２４】次にキャッシュ３をクライアントＢ、Ｃ、Ｄにそれぞれキャッシュ３Ａ、３Ｂ、３Ｃの領域を割り当てている場合に、クライアントＢ、Ｃ、Ｄより分割時の優先順位の高いクライアントＡがデータＡをアクセスする。クライアントＡは他のクライアントと同様の手順にてクライアントＡのＩＤをネットワーク８に送信し、サーバーシステム１のアクセス権利を得る。受信したＩＤがクライアントＡであると認識したＣＰＵ２は、キャッシュ３の使用状況をテーブル５のビットマップ５Ａにより判断する。ＣＰＵ２はビットマップ５Ａの各ビット５Ａ１～５Ａ６が１でキャッシュ３に空きがないと判断すると、テーブル５の分割キャッシュの詳細情報５Ｂからクライアント ＩＤ５Ｂ１を読み取り、キャッシュ３を使用中であるクライアントＢ、Ｃ、Ｄの中で最も分割時の優先順位が低いクライアントＤとアクセス要求元であるクライアントＡのキャッシュ分割時の優先順位をＩＤにより比較する。キャッシュ分割時の優先順位がクライアントＤよりクライアントＡが高いと判断したＣＰＵ２

は、キャッシュ３Ｃ内のデータ７Ｄをメモリ４もしくは、外部記憶装置７に退避後、キャッシュ３Ｃ領域の詳細情報５Ｂをクライアント Ｄからクライアント Ａに書き換え、キャッシュ３Ｃをクライアント Ａ用に割り当てる。割り当てたキャッシュ３Ｃには、外部記憶装置７から読みだしたデータ７Ａを転送、蓄積し、クライアント Ａは、キャッシュ３Ｃ内のデータ７Ａをアクセスする。もちろん、このアクセスにより、クライアント Ｄを除く他のクライアントのキャッシュ３内のデータが上書きされることはない。

【００２５】以上のように、サーバーシステムがファイルアクセス要求を発生させた各クライアントのクライアントＩＤを認識し、各クライアントに対して専用のキャッシュ領域を割り当てることにより、他のクライアントからのファイルアクセスに影響されず、キャッシュのもつ機能を有効に活かしたクライアント／サーバーシステムを提供できる。

【００２６】本実施例においては、キャッシュ領域に空きがなく、サーバーシステムにアクセスしたクライアントがキャッシュを占有しているクライアントと比較して最も優先順位が低い場合、キャッシュを割り当てることなく、直接、外部記憶装置をアクセスすることとしたが、キャッシュを占有しているクライアントのなかで最も優先順位の低いクライアントのキャッシュ領域をアクセスしたクライアントのキャッシュ領域として割り当てを行なってもよいし、各クライアントのサーバーシステムへのアクセス頻度から、キャッシュ領域の割り当てを行なってもよい。

【００２７】キャッシュの分割制御は、サーバーシステム内のＣＰＵに限らず専用のシステムを用いて行なってもよいし、複数のキャッシュを用いて同様の制御を行なってもよい。

【００２８】さらに、本発明を用いればクライアント／サーバー間の高速なデータ処理が行えることから、クライアント側でキャッシュを持つ必要がなくなり、安価なクライアントを用いたクライアント／サーバーシステムを提供できる。

【００２９】本発明をいくつかの例によって説明したが、本発明はこの実施例に限定されるものでない。本発明を使用できるシステム構成は多数あり、本発明の精神及び範囲から逸脱することなしに、形態及び細部に種々の変更を加えることが可能である。

【００３０】

【発明の効果】本発明によれば、ハードディスクなどの外部記憶装置と外部記憶装置から読み出したデータを蓄積するキャッシュを備えたサーバーシステムにおいて、サーバーシステムとファイルアクセス要求元との間のファイルアクセスの処理速度を向上することができる。

【図面の簡単な説明】

【図１】クライアント／サーバーシステムのブロック構成図。

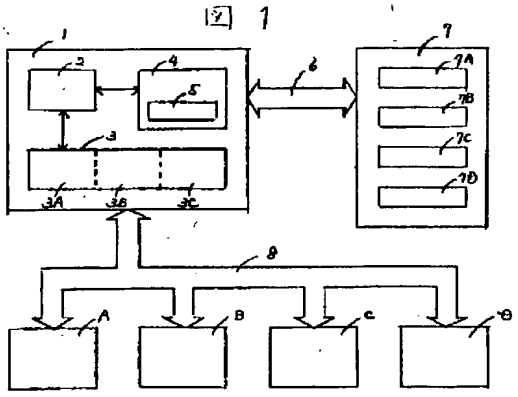
【図２】キャッシュ領域の使用状況を管理するキャッシュアロケーションテーブルを示す図。

【符号の説明】

１…サーバーシステム、２…ＣＰＵ、３…キャッシュ、３Ａ～３Ｃ…分割キャッシュ領域、４…メモリ、５…キャッシュアロケーションテーブル、５Ａ…ビットマップ、５Ａ１，５Ａ２…分割キャッシュ３Ａの使用状態を管理するビット、５Ａ３，５Ａ４…分割キャッシュ３Ｂの使用状態を管理するビット、５Ａ５，５Ａ６…分割キャッシュ３Ｃの使用状態を管理するビット、５Ｂ…分割キャッシュ領域の詳細情報、５Ｂ１…クライアントＩＤ、５Ｂ２…先頭ブロックアドレス、５Ｂ３…最終ブロックアドレス、５Ｂ４……レングス、５Ｂ５…ファイル情報、６…入出力バス、７…外部記憶装置、７Ａ～７Ｄ…データ、８…ネットワーク、Ａ～Ｄ……クライアント

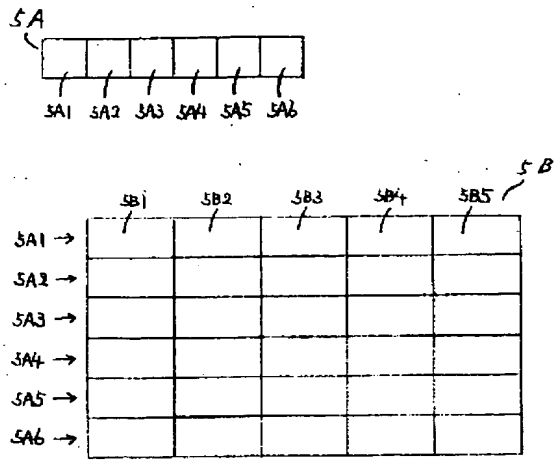
【図１】

図 1



【図２】

図２


5A
5A1
5A2
5A3
5A4
5A5
5A6
5B
5B1
5B2
5B3
5B4
5B5
5A1
5A2
5A3
5A4
5A5
5A6